(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification[7]: C07H 21/00

(21) International Application Number: PCT/GB01/04084

(22) International Filing Date:
12 September 2001 (12.09.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0022330.5   12 September 2000 (12.09.2000)   GB

(71) Applicant *(for all designated States except US)*: SANG-AMO BIOSCIENCES, Inc. [US/US]; Point Richmond Tech Center, 501 Canal Boulevard, Suite A 100, CA 94804 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: CHOO, Yen [GR/GB]; 130 Ross Street, Cambridge CB1 3BU (GB). ISALAN, Mark [GB/GB]; 24 Shottfield Avenue, East Sheen, London SW14 8EA (GB).

(74) Agent: KHOO, Chong-Yee; D Young & Co., 21 New Fetter Lane, London EC4A 1DA (GB).

(81). Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:
—   *of inventorship (Rule 4.17(iv)) for US only*

Published:
—   *with international search report*
—   *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHOD FOR THE PREPARATION OF SELECTIVELY RANDOMISED NUCLEIC ACID MOLECULES

(57) Abstract: We describe methods for producing selectively randomised oligonucleotides using a phosphoramidite dinucleotide and mononucleotide synthesis strategy. These methods comprise: (i) coupling a dinucleotide phosphoramidite to the 3' position of a nucleotide; (ii) coupling a mononucleotide phosphoramidite to the oligonucleotide of (i); and (iii) repeating steps (i) and (ii) until the desired length oligonucleotide is produced. This methodology provides randomised oligonucleotides without the problems of NNN randomisation, without having to resort to complicated resin-splitting procedures or the use of low coupling efficiency trinucleotide phosphoramidites.

METHOD FOR THE PREPARATION OF SELECTIVELY RANDOMISED NUCLEIC ACID MOLECULES

FIELD OF THE INVENTION

The present invention relates to methods for the production of selectively randomised nucleic acid molecules. In particular, the invention relates to the production of

5   population(s) of such nucleic acid molecules, and to the construction of libraries comprising such molecules.

BACKGROUND TO THE INVENTION

In prior art schemes randomised DNA sequences are synthesised by sequentially coupling a mixture of the four nucleoside precursors to the growing oligonucleotide. In

10   this way all 64 possible codon sequences, including the 3 possible stop codons are generated (i.e., NNN).

This strategy has been improved by exploiting the third position redundancy of many codon assignments. By using all four nucleosides in the first two codon positions, but only G and C or A and T in the third position (i.e., NNG/C or NNA/T), it is possible to

15   produce 32 different triplets encoding all 20 amino acids. In this manner, the bias in favour of the amino acids encoded by multiple codon sequences is maintained, and the presence of a stop codon will produce truncated amino acid sequences upon translation. This truncation, which occurs with a frequency of $(n/32)$ where n is the number of amino acids of the randomised sequence, considerably limits the complexity that can be achieved for

20   long randomised peptide libraries. With this strategy, introducing subsets of the 20 amino acids at a given position in the molecule, e.g. to exclude the codon corresponding to the wild type sequence, is limited to only those combinations that can be generated through the synthesis of mixtures of monomers.

Another prior art approach is based on the synthesis of individual codon sequences

25   to the growing oligonucleotide on separate columns, as described for the synthesis of random peptide libraries (Lam et al (1991) Nature 354 pp 82-84). After synthesis of each

codon, the beads from all the columns are mixed together, spilt again and then repacked
into new columns to synthesise the next codon. This resin-splitting method allows
randomisation of a codon position to be modulated by varying the proportions of starting
materials and/or reaction products mixed at each step. However, the benefits are countered
5   by having to use several columns and a procedure that becomes increasingly laborious
with the complexity of the mutagenesis scheme, making the task exceedingly time-
consuming.

Another prior art method is based on the use of 20 pre-synthesised codons as
monomeric units. This method involves the preparation of trinucleotide phosphoramidites
10   and their use in synthesising randomised oligonucleotide sequences by automated DNA
synthesis methodology (e.g. Lyttle et al (1995) Biotechniques 19 pp 274-280; Ono et al
(1995) NAR 23 pp 4677-4682; Virnekäs et al (1994) NAR 22 pp 5600-5607). Although
this method offers control over particular subsets of residues at a given position, the
synthesis and efficient coupling of trinucleotide blocks is not a straightforward process.
15   Attempts to use triplets for the generation of protein mutants are hindered by low coupling
efficiency, as well as deletions in the final product. Such protocols also generate a
significant amount of single-base insertions.

Ono et al. perform synthesis of the antisense codon triplets in the 5'-3' direction.
These anti-codons are then converted into the sense strand by *in vitro* replication methods.
20   In this case too, more study would be required to establish optimal conditions for coupling
reactions to achieve equimolar incorporation of the codons. Synthesis of the 20 triplet
blocks is long and complicated and these triplets are in any case obtained in low yields.

Neuner et al (1995) NAR 26 pp 12233-1227) describe a codon-based mutagenesis
strategy using dinucleotide phosphoramidite building blocks within a resin-splitting
25   framework.

Thus, generating randomised oligonucleotide molecules according to the prior art
involves a high risk of introducing stop codons, and/or the incorporation of many non-

optimal or rare codons into the coding sequences produced in this manner. This can lead to early truncation of expressed sequences, as well as inefficient expression due to ribosome stalling caused by low concentrations of rare tRNAs. Furthermore, generation of randomised oligonucleotides using trinucleotide phosphoramidites is hindered by low

5   coupling efficiencies and difficulties in the synthesis of the trinucleotides themselves. Prior art production of synthetic oligonucleotide molecules using dinucleotides involves resin-splitting techniques which are extremely labour intensive. It is clearly desirable to produce randomised oligonucleotides without the problems of NNN randomisation, without having to resort to complicated resin-splitting procedures or the use of low

10   coupling efficiency trinucleotide phosphoramidites.

SUMMARY OF THE INVENTION

We describe techniques for producing selectively randomised oligonucleotides. Such selectively randomised oligonucleotides may be synthesised for example using an automated nucleic acid synthesiser.

15   In particular, we describe the synthesis of oligonucleotides stepwise by the addition of a dinucleotide phosphoramidite to the growing oligonucleotide chain, followed by the addition of a mononucleotide phosphoramidite to the growing oligonucleotide chain. The oligonucleotide is thus built up to the desired length by repetitions of this synthesis scheme.

20   The method may in particular comprise the steps of: (i) coupling a dinucleotide phosphoramidite to the 3' position of a nucleotide base; and (ii) coupling a mononucleotide phosphoramidite to the oligonucleotide of (i). Optionally, steps (i) and (ii) are repeated until the desired length oligonucleotide is produced. The oligonucleotide may be synthesised on a solid support.

25

'Selectively randomised' means that the scheme is designed to allow randomisation to be limited to particular subsets of sequences encoding different subsets of amino acids/STOP codons as desired by the operator. Further, the selectivity applies to the choice of position(s) within the oligonucleotide which are randomised. For example,

5 according to the methods described here it is possible to retain a certain fixed sequence framework, whilst incorporating particular directed randomisation(s) therein.

Selectivity is introduced into the scheme of synthesis according to the alternative amino acids which it is desired to encode at the relevant positions in the growing oligonucleotide. The amino acids for a particular position are first chosen. This/these may

10 be a single (specified) amino acid such as ALA, or may be the entire pool of twenty amino acids plus the possibility of a STOP codon, or any intermediate combination. Particular pools or subsets of such codon(s) may be chosen, according to the design of the oligonucleotide which they desire to produce. For each amino acid which it is desired to include in the pool of codon randomisation, the relevant dinucleotide phosphoramidite is

15 selected, for example by reference to Table 1. A cocktail or mixture of such phosphoramidites is then used to extend the oligonucleotide, thereby providing a selected pool of possible randomisation(s) which will be introduced into the oligonucleotide as explained herein. Examples of this selective randomisation approach are discussed in more detail below.

20 We describe a method for making a selectively randomised synthetic oligonucleotide comprising providing a starting material coupled to a suitable support in a nucleic acid synthesiser; deprotecting said starting material at the 3' position; coupling a dinucleotide phosphoramidite to said 3' position; deprotecting the new 3' position of the extended oligonucleotide chain; coupling a mononucleotide phosphoramidite to said 3'

25 position, and repeating the two sets of deprotecting/coupling steps until the desired length oligonucleotide is produced.

5

The term 'starting material' means any suitable solid phase support for the synthesis of oligonucleotides thereon, such as commercially available resins for use in conjunction with an Applied Biosystems automated DNA/RNA synthesiser.

Furthermore, we disclose a method as described above, wherein at each repetition
5   of the dinucleotide coupling step a specific dinucleotide phosphoramidite or specific mixture of dinucleotide phosphoramidite(s) is used in the coupling. This specificity is introduced according to the choices of the person working the methods as mentioned above and explained in detail below.

Preferably, at each repetition of the dinucleotide coupling step, a dinucleotide
10   phosphoramidite or mixture of dinucleotide phosphoramidite(s) selected from Table 1 is incorporated into the growing oligonucleotide molecule. The dinucleotide phosphoramidites may be selected from the group consisting of: AA, AA, AT, AT, CA, CA, GA, GA, TG, TG, AC, CC, CG, CT, GC, GG, GT, TC, TA, and TT.

Further details on the selection of phosphoramidites or pools thereof may be found
15   below.

Preferably, a mixture of mononucleotide phosphoramidites is employed. Preferably, said mononucleotide phosphoramidite comprises a mixture of G/C mononucleotide phosphoramidites. Preferably, said mononucleotide phosphoramidite comprises a 50:50 ratio of G:C mononucleotide phosphoramidites. Preferably, said
20   mononucleotide phosphoramidite comprises a mixture of A/T mononucleotide phosphoramidites. Preferably, said mononucleotide phosphoramidite comprises a 50:50 ratio of A:T mononucleotide phosphoramidites.

Preferably, said oligonucleotide comprises an open reading frame (ORF). An open reading frame means a stretch of nucleic acid which comprises a series of codons capable
25   of being translated into a polypeptide by the appropriate cellular transcription/translation

machinery. In practice, the term ORF refers to any such series of codons up to the first occurring STOP codon (i.e. TAG, TAA or TGA).

In a preferred embodiment, the ORF comprises an optimal human codon. The optimal human codon may be selected from the optimal human codons shown below in Table 1. Preferably, most or substantially all of the codons in the ORF comprise optimal human codons, preferably as shown below in Table 1.

Preferably, codons of said ORF are synthesised from dinucleotide phosphoramidites (XY) and mononucleotide phosphoramidites (Z) in the XY-Z conformation such that the dinucleotide phosphoramidite forms the first two bases of the codon, the mono nucleotide phosphoramidite forming the last base of the codon.

Preferably, the oligonucleotide encodes a zinc finger polypeptide; preferably, said oligonucleotide ORFs comprise nucleotide sequence(s) encoding one or more zinc finger motif(s) or part(s) thereof. Oligonucleotides as described here are preferably oligonucleotides encoding a zinc finger motif. A zinc finger is a DNA-binding protein domain that may be used as a scaffold to design DNA-binding proteins. Preferably, oligonucleotides as described here are oligonucleotides encoding a zinc finger nucleic acid binding motif. The properties of such motifs include the possession of a Cys2-His2 motif, and are discussed in more detail below. Thus, we further disclose a method as described above, wherein said oligonucleotide ORFs are selectively randomised at positions other than the conserved CYS-HIS motif(s).

In a preferred embodiment of the invention, the oligonucleotide is selectively randomised at a base contacting position of the zinc finger polypeptide. More preferably, the oligonucleotide is partially randomised; that is, the oligonucleotide is randomised at one or more base contacting positions, while the remaining base contacting position or positions are fixed. However, the oligonucleotide may be selectively randomised at substantially all the base contacting positions of the zinc finger polypeptide.

Oligonucleotides described here, or made by the methods described here, may also be used as targets or substrates for zinc finger binding. Furthermore, they may be used to construct libraries of nucleic acid targets for use in selection or screening, for example, for selection of nucleic acids capable of binding a particular zinc finger sequence.

5   In another aspect, the invention relates to oligonucleotides produced by a method as described above.

In another aspect, the invention relates to a library comprising oligonucleotides produced by a method as described above.

In another aspect, the invention relates to a library as described above wherein said

10   library is a phage display library.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a diagram and a selective randomisation scheme: construction of a gene cassette coding for a zinc finger phage display library with 'smart' randomisations. (A) A scheme for generation of selective randomisation throughout the α-helix of a zinc

15   finger. A set of complementary oligonucleotides is used to construct a series of "minicassettes" which can be annealed and ligated together to construct the randomised portion of the gene. After ligation of all the minicassettes, the full-length construct is recovered by PCR using primers which contain *Sfi*I/*Not*I restriction enzyme sites for cloning into phage vector. (B) Examples of the oligonucleotides used to achieve selective

20   randomisation of a zinc finger protein.

DETAILED DESCRIPTION OF THE INVENTION

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of chemistry, molecular biology, microbiology, recombinant DNA and immunology, which are within the capabilities of a person of ordinary skill in

the art. Such techniques are explained in the literature. See, for example, J. Sambrook, E.

F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual*, Second

Edition, Books 1-3, Cold Spring Harbor Laboratory Press; Ausubel, F. M. et al. (1995 and

periodic supplements; *Current Protocols in Molecular Biology*, ch. 9, 13, and 16, John

5   Wiley & Sons, New York, N.Y.); B. Roe, J. Crabtree, and A. Kahn, 1996, *DNA Isolation*

and Sequencing: Essential Techniques, John Wiley & Sons; J. M. Polak and James O'D.

McGee, 1990, *In Situ Hybridization: Principles and Practice*; Oxford University Press; M.

J. Gait (Editor), 1984, *Oligonucleotide Synthesis: A Practical Approach*, Irl Press; and, D.

M. J. Lilley and J. E. Dahlberg, 1992, *Methods of Enzymology: DNA Structure Part A:*

10   *Synthesis and Physical Analysis of DNA* Methods in Enzymology, Academic Press. Each

of these general texts is herein incorporated by reference.

SELECTIVE RANDOMISATION

Randomisation according to the methods described here is selective. This means

that the degree of randomisation applied to each position within the oligonucleotide being

15   synthesised is selected by the person working the method.

This is different to prior art techniques such as NNN randomisation which method

produces oligonucleotide(s) which are completely (i.e., non-selectively) randomised.

The degree of randomisation at each position refers to whether a particular amino

acid is specified by supplying only the appropriate single dinucleotide phosphoramidite (or

20   the appropriate sequence of three mononucleotide phosphoramidites) at that stage in the

synthesis (i.e., a specified position), or whether one of only a few amino acids is specified

by supplying only one (eg. TA or TT) or a few different dinucleotide phosphoramidite(s)

at that stage in the synthesis (i.e., a low degree of randomisation position), or whether

potentially any amino acid is specified by supplying all 15 dinucleotide phosphoramidites

25   as shown in Table 1, which may result in potentially any amino acid being specified (i.e., a

high degree of randomisation or complete randomisation).

This selectivity enables the oligonucleotide which is being manufactured to be designed with a certain degree of randomisation at each position within the oligonucleotide. The degree of randomisation is chosen (i.e., selected) by the person working the method, and is advantageously brought about using the methods as explained

5   herein.

This selectivity is accomplished by supplying different combinations of dinucleotide phophoramidites to the extension mix at each round of oligonucleotide extension. To incorporate a completely randomised position within the growing oligonucleotide, a combination of all 15 dinucleotide phosphoramidites may be supplied to

10   the extension mix, followed by a 50:50 mix of G:C mononucleotide phosphoramidites. This results in an oligonucleotide being generated which may encode any amino acid at that particular codon, or may even terminate at that particular codon (i.e., a stop codon could be introduced at that position as could any other coding codon).

It is an advantage that certain position(s) within the oligonucleotide may be

15   effectively specified by supplying a single dinucleotide phosphoramidite to the growing oligonucleotide at the appropriate point in the process of synthesising said oligonucleotide. Clearly, such positions may equally be specified by supplying the relevant three mononucleotides to the growing oligonucleotide at the appropriate point, according to convenience or preference of the operator.

20   Further, it is an advantage that codons may be selectively randomised. This means that the codon may be randomised across a subset of particular amino acid residues, rather that across the whole spectrum of possible amino acids. Thus, selective randomisation refers to the selection of a subset of possible amino acids which it is desired to randomly introduce at a given position, and of the corresponding codons encoding them. These

25   codons have corresponding dinucleotide phosphoramidites which may be found by reference to Table 1 herein. The oligonucleotide may then be extended by supplying a subset of dinucleotides which give rise to the optimal human codons for the particular amino acids which it is desirable to randomly incorporate at said position in said

10

oligonucleotide. Each of the different dinucleotides supplied is thus incorporated with similar probability, leading to a selective randomisation of said position in the oligonucleotide wherein the codon incorporated into the nucleic acid molecule at that point is randomly determined, but from a subset of selected possible codons rather from
5   the full spectrum of all codons. Thus, the restricted or selective randomisation is effected.

For a fully randomised position within the oligonucleotide, all 15 dinucleotides are introduced into the mixture at the appropriate point in the synthesis, followed by an equal mixture of G and C mononucleotides at the subsequent step.

For a selectively randomised position within the oligonucleotide, fewer than 15
10   dinucleotides are introduced into the mixture. For example, to introduce a selective randomisation between codons encoding LYS or ASN or MET or ILE or GLN or HIS or GLU or ASP or TRP or CYS, a subset of the 15 dinucleotides would be introduced, followed by the appropriate G/C mixture. the particular subset of dinucleotides which would be introduced may be determined with reference to Table 1. In this example, the
15   appropriate subset of dinucleotides to introduce would be AA, AT, CA, GA and TG, followed by an equal mixture of G and C mononucleotides in the subsequent step.

Table 1

Nucleotide		Optimal		Non-optimal	
Di-	Mono	human codon	(codes for)	codon	(codes for)
XY	Z	XY-Z	(aa)	XY-Z	(aa)
AA	G	AA-G	Lys	N/A	
AA	C	AA-C	Asn	N/A	
AT	G	AT-G	Met	N/A	
AT	C	AT-C	Ile	N/A	
CA	G	CA-G	Gln	N/A	
CA	C	CA-C	His	N/A	
GA	G	GA-G	Glu	N/A	
GA	C	GA-C	Asp	N/A	
TG	G	TG-G	Trp	N/A	
TG	C	TG-C	Cys	N/A	
AC	C	AC-C	Thr	AC-G	Thr
CC	C	CC-C	Pro	CC-G	Pro
CG	C	CG-C	Arg	CG-G	Arg
CT	C	CT-C	Leu	CT-G	Leu
GC	C	GC-C	Ala	GC-G	Ala
GG	C	GG-C	Gly	GG-G	Gly
GT	C	GT-C	Val	GT-G	Val
TC	C	TC-C	Ser	TC-G	Ser
TA	C	TA-C	Tyr	TA-G	STOP
TT	C	TT-C	Phe	TT-G	Leu

Dinucleotides are shown in italics in Column 1; mononucleotides are shown underlined in Column 2. Corresponding dinucleotide phosphoramidites and mononucleotide phosphoramidites may be synthesised.

12

A selectively randomised position is a position at which codons are randomly incorporated from a set of possible codons which could be incorporated which includes from one to twenty possible coding codons, and may include one or more stop codons. The selective part of the randomisation occurs by choosing the set of possible codons

5 which may be incorporated therein. This set of codons may be less than the full range of possible codons. A selectively randomised position so produced may be directed, for example directed towards codons specifying aromatic residues, or it may be excluded, for example any codon except those encoding SER, or it may be specific, for example any codon encoding ALA (or a subset of codons encoding ALA), or may be any combination

10 of such strategies or any other desirable strategy. Our methods allow oligonucleotides to be conveniently synthesised according to the strategy desired by the person working the method, whilst facilitating the production of oligonucleotides comprising codons, said codons being randomly chosen from varying pools of possible codons, said pools being chosen by the person working the method, such as with reference to Table 1 and the

15 methods described herein.

For example, to synthesise a directed position, such as directed towards acidic residues ASP or GLU, at the appropriate step in the synthesis a pool of dinucleotide phosphoramidites comprising the GA dinucleotide is supplied to the growing oligonucleotide chain, followed by a solution of a 50:50 mixture of G:C mononucleotide

20 phosphoramidites at the subsequent step. In this way, 100% of the oligonucleotides produced should have a codon specifying an acidic residue at said position, 50% specifying GLU (i.e., GAG), 50% specifying ASP (GAC).

For example, to synthesise a directed position, such as directed towards phosphoacceptor residues SER, THR, TYR at the appropriate step in the synthesis a pool

25 of dinucleotide phosphoramidites comprising TC, AC, TA is supplied to the growing oligonucleotide chain, followed by a solution of C mononucleotide phosphoramidite at the subsequent step. In this way, 100% of the oligonucleotides produced will have a codon specifying a phosphoacceptor residue at said position, whilst avoiding the possibility of incorporating a STOP codon at said position. To illustrate how straightforward the

30 methods described here are to adapt to different combinations, by replacing the C

mononucleotide with a mixture of G:C mononucleotides, the person working the method may introduce the possibility that a STOP codon will be incorporated into the oligonucleotide, whilst retaining the specificity of introducing a phosphoacceptor codon. If 100% C is used in the second step, the three possible codons introduced are TCC (SER),

5   ACC (THR), or TAC (TYR). If a mixture of G and C is used in the second step, the six possible codons introduced are TCC, TCG (SER); ACC, ACG (THR); TAC (TYR) and TAG (STOP). By altering the proportion of C:G in the mixture, the likelihood of introducing a STOP codon may easily be manipulated, without compromising the possibility of introducing a sequence encoding a non-phosphoacceptor residue into the

10   oligonucleotide, demonstrating the versatility of the methods described here.

For example, to synthesise a directed position, such as directed towards aromatic residues PHE, TYR, TRP, at the appropriate step in the synthesis a pool of dinucleotide phosphoramidites comprising TG, TA, TT is supplied to the growing oligonucleotide chain, followed by a solution of a 50:50 mixture of G:C mononucleotide phosphoramidites

15   at the subsequent step. In this way, 50% of the oligonucleotides produced will have a codon specifying an aromatic residue at said position.

For example, to synthesise a directed position, such as directed towards aliphatic residues GLY, ALA, VAL, LEU, ILE, at the appropriate step in the synthesis a pool of dinucleotide phosphoramidites comprising AT, CT, GC, GG, GT is supplied to the

20   growing oligonucleotide chain, followed by a solution of C mononucleotide phosphoramidite at the subsequent step. In this way, 100% of the oligonucleotides produced will have a codon specifying an aliphatic residue at said position.

For example, to synthesise an excluded position, such as any amino acid except PRO, at the appropriate step in the synthesis a pool of dinucleotide phosphoramidites ·

25   comprising all dinucleotides shown in Table 1 except CC (i.e., comprising AA, AT, CA, GA, TG, AC, CG, CT, GC, GG, GT, TC, TA and TT) is supplied to the growing oligonucleotide chain, followed by a solution of a 50:50 mixture of G:C mononucleotide phosphoramidites at the subsequent step. In this way, none of the oligonucleotides

produced will have a codon specifying a proline residue at said position. Clearly, it will not be necessary to include all 14 dinucleotides unless it is wished that codons specifying all amino acids bar PRO are included in the oligonucleotide at this position. Excluding a particular amino acid such as PRO may be combined with any other strategy simply and conveniently by altering the composition of the dinucleotide mixture appropriately, for example with reference to Table 1.

For example, to synthesise a directed position, such as directed towards such as non-phosphorylatable residues LYS, ASN, MET, ILE, GLN, HIS, GLU, ASP, TRP, CYS, PRO, ARG, LEU, ALA, GLY, VAL, PHE, at the appropriate step in the synthesis a pool of dinucleotide phosphoramidites comprising AA, AT, CA, GA, TG, CC, CG, CT, GC, GG, GT, TT, is supplied to the growing oligonucleotide chain, followed by a solution of a 50:50 mixture of G:C mononucleotide phosphoramidites at the subsequent step. In this way, none of the oligonucleotides produced should have a codon specifying a phosphorylatable residue at said position.

For example, to synthesise a specific position, such as SER, at the appropriate step in the synthesis a pool of dinucleotide phosphoramidites comprising TC is supplied to the growing oligonucleotide chain, followed by a solution of a 50:50 mixture of G:C mononucleotide phosphoramidites at the subsequent step. In this way, 100% of the oligonucleotides produced will have a codon specifying an SER at said position. Clearly, the same result would be achieved if a solution of G mononucleotide phosphoramidite or a solution of C mononucleotide phosphoramidite replaced the solution of a 50:50 mixture of G:C mononucleotide phosphoramidites in this example, since both codons TCC and TCG specify SER. This might be desirable, for example to save labour when producing said oligonucleotide(s), which is a further advantage of the versatile method described here.

Any other directed approach may be designed and implemented, for example by picking the appropriate combination(s) of dinucleotide(s) from Table 1 and proceeding with the synthesis as taught herein.

Useful groups of chemically related amino acids which are often regarded as similar when performing mutagenesis/randomisation as discussed herein are found in the following table. Conserved substitutions may be made according to the following table which indicates some possible conservative substitutions, where amino acids on the same

5   block in the second column and preferably in the same line in the third column may be substituted for each other. For some instances other conserved substitutions may be made.

ALIPHATIC	Non-polar	G A P
		I L V
	Polar - uncharged	C S T M
		N Q
	Polar - charged	D E
		K R
AROMATIC		H F W Y
OTHER		N Q D E

Further, amino acids may be grouped together according to their properties eg. ASP, GLU both have acidic side chains, LYS, ARG, HIS each have basic side chains, ASN and GLN both have amide side chains, CYS and MET have sulphur-containing side

10   chains, PHE, TYR and TRP have aromatic side chains, SER and THR both have aliphatic hydroxyl side chains, PRO has a secondary amino group, GLY, ALA, VAL, LEU and ILE all have aliphatic (or small neutral) side chains.

Clearly, variations on the scheme presented in Table 1 will fall within the scope of methods disclosed here. For example, if a different set of optimised codons were desired,

15   such as for optimal expression in a different organism, then the methods are easily adapted to produce oligonucleotides encoding same by compiling a table using the alternative codon optimisation data for the target organism. An example of such a target organism is yeast. For example, optimal yeast codons can be achieved by adding the mononucleotides A, or T, (or a 50:50 mixture of A:T mononucleotides) after any or all, of the 16 possible

20   optimal yeast dinucleotides. These include GC-T=ALA; AG-A=ARG; AA-T=ASN; AT-A=ILE; GA-T=ASP; TG-T=CYS; CA-A=GLN; GA-A=GLU; GG-T=GLY;

CA-T=HIS; TT-A=LEU; AA-A=LYS; CC-A=PRO; TT-T=PHE; TC-T=SER;
AC-T=THR; TA-T=TYR; GT-T=VAL. Possible variants of this scheme are TG-G=TRP
and AT-G=MET, which are optimal yeast codons having G mononucleotides at the 3'
position of said codons. The methods may be advantageously adapted in this fashion for
5   any other desired target organism using the appropriate codon optimisation data.

MAKING OF OLIGONUCLEOTIDES

Synthetic oligonucleotide(s) may be synthesised in the following manner: a
40nmol scale column is provided on a DNA synthesiser such as an Applied Biosystems
DNA/RNA synthesiser.

10   Bases are coupled according to standard phosphoramidite chemistry; the 3' end of
the growing oligonucleotide chain is deprotected; a mixture of dinucleotide
phosphoramidites is applied to the column, said mixture being determined according to the
pool of possible codons which it is wished to incorporate, for example by reference to
Table 1 by working from (i.e., back-translating) the pool of amino acid(s) which it is
15   desired to incorporate into the ORF.

The coupling reaction is then performed, which may be extended to two minutes or
even more, the 3' end of the growing oligonucleotide chain is deprotected again, and a
G/C coupling is performed, which may be any proportion of G:C including 100% G or
100% C according to the needs of the operator, for example with reference to Table 1.
20   Typically, the mixture will be 50% G: 50% C.

The deprotecting and coupling of dinucleotide, followed by the deprotecting and
coupling of mononucleotide steps are repeated in sequence until the desired length of
oligonucleotide is produced, varying the composition of the dinucleotide and/or
mononucleotide mixture(s) according to the desired composition of the resulting
25   oligonucleotide at the various positions being synthesised.

PURIFICATION OF OLIGONUCLEOTIDES

Purification or 'cleaning up' of oligonucleotides may be accomplished by any suitable means known in the art. These include HPLC chromatographic purification, ethanol precipitation, preparative polyacrylamide gel electrophoresis, use of commercially

5   available spin-columns or any other suitable purification means known in the art.

Oligonucleotides may be of any length. If technical limits of synthesis constrain the desired length of the oligonucleotide, then more than one oligonucleotide may be conjoined, such as via ligation, to produce longer oligonucleotide(s). Nucleic acids may be engineered using recombinant DNA techniques discussed below to comprise one or more

10   oligonucleotides, in which case these nucleic acids may be referred to herein as oligonucleotides by virtue of the fact that they comprise same.

ZINC FINGERS

A zinc finger is a DNA-binding protein domain that may be used as a scaffold to design DNA-binding proteins with predetermined sequence-specificity (Klug, A. &

15   Rhodes, D. (1987) 'Zinc fingers': a novel protein motif for nucleic acid recognition. *Trends Biochem. Sci.* **12**, 464-469; Choo, Y. & Klug, A. (1995) Designing DNA-binding proteins on the surface of filamentous phage. *Curr. Opin. Biotech.* **6**, 431-436). The peptide motif comprises about 30 amino acids that adopt a compact DNA-binding structure on chelating a zinc ion (Miller, J., McLachlan, A. D. & Klug, A. (1985) Repetitive zinc-binding

20   domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J* **4**, 1609-1614). Each zinc finger module is capable of recognising 3-4bp of DNA, such that arrays comprising tandemly repeated modules bind proportionally longer nucleotide sequences. The crystal structure of the Zif268 DNA-binding domain, in complex with its optimal DNA binding site, shows that the zinc finger array wraps around the DNA, with the α-

25   helix of each finger buried in the major groove (Pavletich, N. P. & Pabo, C. O. (1991) Zinc finger-DNA recognition: Crystal structure of a Zif268-DNA complex at 2.1 Å. *Science* **252**, 809-817).

The geometrical properties of zinc finger structures mean that a versatile binding surface can be created by varying a small number of amino acid positions on each finger's central α-helix. Moreover, zinc fingers may be linked together to bind to longer, contiguous stretches of DNA. Large randomised libraries of zinc fingers have been

5   engineered by phage display, so that zinc finger variants are displayed on the viral capsid. Such libraries have been extensively screened to select fingers that bind to various duplex DNA sequences (Choo, Y., & Klug, A. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 11163-11167. Greisman, H. A., & Pabo, C. O. (1997) *Science 275*, 657-661. Jamieson, A. C., Kim, S.-H., & Wells, J. A. (1994) *Biochemistry 33*, 5689-5695. Wu, H., Yang, W.-P., &

10   Barbas III, C. F. (1995) *Proc. Natl. Acad. Sci. USA 92*, 344-348. Isalan, M., Klug, A., & Choo, Y. (1998) *Biochemistry 37*, 12026-12033.)and to RNA (Friesen, W. J., & Darby, M. K. (1997) *J Biol Chem 272*, 10994-10997. Friesen, W. J., & Darby, M. K. (1998) *Nat Struct Biol 5*, 543-546. Blancafort, P., Steinberg, S. V., Paquin, B., Klinck, R., Scott, J. K., & Cedergren, R. (1999) *Chemistry and Biology 6*, 585-597.).

15   Zinc fingers, as is known in the art, are nucleic acid binding molecules. A zinc finger binding motif is a structure well known to those in the art and defined in, for example, Miller *et al.*, (1985) EMBO J. 4:1609-1614; Berg (1988) PNAS (USA) 85:99-102; Lee *et al.*, (1989) Science 245:635-637; see International patent applications WO 96/06166 and WO 96/32475, corresponding to USSN 08/422,107, incorporated herein by

20   reference.

As used herein, "nucleic acid" refers to both RNA and DNA, constructed from natural nucleic acid bases or synthetic bases, or mixtures thereof.

All of the nucleic acid-binding residue positions of zinc fingers, as referred to herein, are numbered from the first residue in the α-helix of the finger, ranging from +1 to

25   +9. "-1" refers to the residue in the framework structure immediately preceding the α-helix in a Cys2-His2 zinc finger polypeptide. Cys2-His2 zinc finger binding proteins, as is well known in the art, bind to target nucleic acid sequences via α-helical zinc metal atom co-ordinated binding motifs known as zinc fingers.

●

19

These and other considerations may be incorporated into a library set of oligonucleotides.

VECTORS

The oligonucleotides can be incorporated into vectors for further manipulation. As used herein, vector (or plasmid) refers to discrete elements that are used to introduce heterologous nucleic acid into cells for either expression or replication thereof. Selection and use of such vehicles are well within the skill of the person of ordinary skill in the art. Many vectors are available, and selection of appropriate vector will depend on the intended use of the vector, i.e. whether it is to be used for DNA amplification or for nucleic acid expression, the size of the DNA to be inserted into the vector, and the host cell to be transformed with the vector. Each vector contains various components depending on its function (amplification of DNA or expression of DNA) and the host cell for which it is compatible. The vector components generally include, but are not limited to, one or more of the following: an origin of replication, one or more marker genes, an enhancer element, a promoter, a transcription termination sequence and a signal sequence.

Both expression and cloning vectors generally contain nucleic acid sequence that enable the vector to replicate in one or more selected host cells. Typically in cloning vectors, this sequence is one that enables the vector to replicate independently of the host chromosomal DNA, and includes origins of replication or autonomously replicating sequences. Such sequences are well known for a variety of bacteria, yeast and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the 2μ plasmid origin is suitable for yeast, and various viral origins (e.g. SV 40, polyoma, adenovirus) are useful for cloning vectors in mammalian cells. Generally, the origin of replication component is not needed for mammalian expression vectors unless these are used in mammalian cells competent for high level DNA replication, such as COS cells.

20

Most expression vectors are shuttle vectors, i.e. they are capable of replication in at least one class of organisms but can be transfected into another class of organisms for expression. For example, a vector is cloned in *E. coli* and then the same vector is transfected into yeast or mammalian cells even though it is not capable of replicating

5   independently of the host cell chromosome. DNA may also be replicated by insertion into the host genome. However, the recovery of genomic DNA is more complex than that of exogenously replicated vector because restriction enzyme digestion is required to excise fragment(s) of such DNA. DNA can be amplified by PCR and be directly transfected into the host cells without any replication component.

10   SELECTABLE MARKERS

Advantageously, an expression and cloning vector may contain a selection gene also referred to as selectable marker. This gene encodes a protein necessary for the survival or growth of transformed host cells grown in a selective culture medium. Host cells not transformed with the vector containing the selection gene will not survive in the

15   culture medium. Typical selection genes encode proteins that confer resistance to antibiotics and other toxins, e.g. ampicillin, neomycin, methotrexate or tetracycline, complement auxotrophic deficiencies, or supply critical nutrients not available from complex media.

As to a selective gene marker appropriate for yeast, any marker gene can be used

20   which facilitates the selection for transformants due to the phenotypic expression of the marker gene. Suitable markers for yeast are, for example, those conferring resistance to antibiotics G418, hygromycin or bleomycin, or provide for prototrophy in an auxotrophic yeast mutant, for example the URA3, LEU2, LYS2, TRP1, or HIS3 gene.

Since the replication of vectors is conveniently done in *E. coli*, an *E. coli* genetic

25   marker and an *E. coli* origin of replication are advantageously included. These can be obtained from *E. coli* plasmids, such as pBR322, Bluescript© vector or a pUC plasmid,

e.g. pUC18 or pUC19, which contain both *E. coli* replication origin and *E. coli* genetic marker conferring resistance to antibiotics, such as ampicillin.

Suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up nucleic acid, such as dihydrofolate reductase (DHFR, methotrexate resistance), thymidine kinase, or genes conferring resistance to G418 or hygromycin. The mammalian cell transformants are placed under selection pressure which only those transformants which have taken up and are expressing the marker are uniquely adapted to survive. In the case of a DHFR or glutamine synthase (GS) marker, selection pressure can be imposed by culturing the transformants under conditions in which the pressure is progressively increased, thereby leading to amplification (at its chromosomal integration site) of both the selection gene and the linked DNA. Amplification is the process by which genes in greater demand for the production of a protein critical for growth, together with closely associated genes which may encode a desired protein, are reiterated in tandem within the chromosomes of recombinant cells. Increased quantities of desired protein are usually synthesised from thus amplified DNA.

EXPRESSION

Expression and cloning vectors usually contain a promoter that is recognised by the host organism and is operably linked to the oligonucleotide(s). Such a promoter may be inducible or constitutive. The promoters are operably linked to the oligonucleotide(s) by removing the promoter from the source DNA by restriction enzyme digestion and inserting the isolated promoter sequence into the vector. Both native promoter sequence(s) (such as nucleic acid binding protein promoter sequences) and many heterologous promoters may be used to direct amplification and/or expression of DNA comprising oligonucleotide(s).

Promoters suitable for use with prokaryotic hosts include, for example, the β-lactamase and lactose promoter systems, alkaline phosphatase, the tryptophan (Trp) promoter system and hybrid promoters such as the tac promoter. Their nucleotide

sequences have been published, thereby enabling the skilled worker operably to ligate them to oligonucleotide(s) as described herein, using linkers or adapters to supply any required restriction sites. Promoters for use in bacterial systems will also generally contain a Shine-Delgarno sequence operably linked to the oligonucleotide(s).

5   Preferred expression vectors are bacterial expression vectors which comprise a promoter of a bacteriophage such as phagex or T7 which is capable of functioning in the bacteria. In one of the most widely used expression systems, the nucleic acid encoding the fusion protein may be transcribed from the vector by T7 RNA polymerase (Studier et al, Methods in Enzymol. 185; 60-89, 1990). In the *E. coli* BL21(DE3) host strain, used in

10   conjunction with pET vectors, the T7 RNA polymerase is produced from the λ-lysogen DE3 in the host bacterium, and its expression is under the control of the IPTG inducible lac UV5 promoter. This system has been employed successfully for over-production of many proteins. Alternatively the polymerase gene may be introduced on a lambda phage by infection with an int- phage such as the CE6 phage which is commercially available

15   (Novagen, Madison, USA). other vectors include vectors containing the lambda PL promoter such as PLEX (Invitrogen, NL) , vectors containing the trc promoters such as pTrcHisXpressTm (Invitrogen) or pTrc99 (Pharmacia Biotech, SE) or vectors containing the tac promoter such as pKK223-3 (Pharmacia Biotech) or PMAL (New England Biolabs, MA, USA).

20   Moreover, the oligonucleotide preferably includes or is conjoined to a secretion sequence in order to facilitate secretion of the encoded polypeptide from bacterial hosts, such that it will be produced as a soluble native peptide rather than in an inclusion body. The peptide may be recovered from the bacterial periplasmic space, or the culture medium, as appropriate. A "leader" peptide may be added to the N-terminal finger.

25   Preferably, the leader peptide is MAEEKP.

Suitable promoting sequences for use with yeast hosts may be regulated or constitutive and are preferably derived from a highly expressed yeast gene, especially a *Saccharomyces cerevisiae* gene. Thus, the promoter of the TRP1 gene, the ADHI or

ADHII gene, the acid phosphatase (PHO5) gene, a promoter of the yeast mating pheromone genes coding for the a- or α-factor or a promoter derived from a gene encoding a glycolytic enzyme such as the promoter of the enolase, glyceraldehyde-3-phosphate dehydrogenase (GAP), 3-phospho glycerate kinase (PGK), hexokinase,

5   pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triose phosphate isomerase, phosphoglucose isomerase or glucokinase genes, or a promoter from the TATA binding protein (TBP) gene can be used. Furthermore, it is possible to use hybrid promoters comprising upstream activation sequences (UAS) of one yeast gene and downstream promoter elements

10   including a functional TATA box of another yeast gene, for example a hybrid promoter including the UAS(s) of the yeast PHO5 gene and downstream promoter elements including a functional TATA box of the yeast GAP gene (PHO5-GAP hybrid promoter). A suitable constitutive PHO5 promoter is e.g. a shortened acid phosphatase PHO5 promoter devoid of the upstream regulatory elements (UAS) such as the PHO5 (-173) promoter

15   element starting at nucleotide -173 and ending at nucleotide -9 of the PHO5 gene.

Oligonucleotide transcription from vectors in mammalian hosts may be controlled by promoters derived from the genomes of viruses such as polyoma virus, adenovirus, fowlpox virus, bovine papilloma virus, avian sarcoma virus, cytomegalovirus (CMV), a retrovirus and Simian Virus 40 (SV40), from heterologous mammalian promoters such as

20   the actin promoter or a very strong promoter, e.g. a ribosomal protein promoter, and from the promoter normally associated with e.g. anucleic acid binding protein sequence, provided such promoters are compatible with the host cell systems.

Transcription of an oligonucleotide by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are relatively orientation and

25   position independent. Many enhancer sequences are known from mammalian genes (e.g. elastase and globin). However, typically one will employ an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270) and the CMV early promoter enhancer. The enhancer may be spliced into the

vector at a position 5' or 3' to the oligonucleotide(s), but is preferably located at a site 5' from the promoter.

Advantageously, a eukaryotic expression vector comprising an oligonucleotide may comprise a locus control region (LCR). LCRs are capable of directing high-level integration site independent expression of transgenes integrated into host cell chromatin, which is of importance especially where the oligonucleotide is to be expressed in the context of a permanently-transfected eukaryotic cell line in which chromosomal integration of the vector has occurred, or in transgenic animals.

Eukaryotic vectors may also contain sequences necessary for the termination of transcription and for stabilising the mRNA. Such sequences are commonly available from the 5' and 3' untranslated regions of eukaryotic or viral DNAs or cDNAs. These regions contain nucleotide segments transcribed as polyadenylated fragments in the untranslated portion of the mRNA produced from the oligonucleotide template.

An expression vector includes any vector capable of expressing oligonucleotide(s) that are operatively linked with regulatory sequences, such as promoter regions, that are capable of expression of such DNAs. Thus, an expression vector refers to a recombinant DNA or RNA construct, such as a plasmid, a phage, recombinant virus or other vector, that upon introduction into an appropriate host cell, results in expression of the cloned DNA. Appropriate expression vectors are well known to those with ordinary skill in the art and include those that are replicable in eukaryotic and/or prokaryotic cells and those that remain episomal or those which integrate into the host cell genome. For example, DNAs encoding nucleic acid binding protein may be inserted into a vector suitable for expression of cDNAs in mammalian cells, e.g. a CMV enhancer-based vector such as pEVRF (Matthias, et al., (1989) NAR 17, 6418).

Particularly useful are expression vectors that provide for the transient expression of the oligonucleotide in mammalian cells. Transient expression usually involves the use of an expression vector that is able to replicate efficiently in a host cell, such that the host

●

cell accumulates many copies of the expression vector, and, in turn, synthesises high levels of protein(s) encoded by the oligonucleotide(s). Transient expression systems are useful e.g. for identifying nucleic acid binding protein mutants, to identify potential phosphorylation sites, or to characterise functional domains of the protein.

5   Construction of vectors employs conventional ligation techniques. Isolated plasmids or DNA fragments are cleaved, tailored, and religated in the form desired to generate the plasmids required. If desired, analysis to confirm correct sequences in the constructed plasmids is performed in a known fashion. Suitable methods for constructing expression vectors, preparing in vitro transcripts, introducing DNA into host cells, and

10   performing analyses for assessing oligonucleotide expression and function are known to those skilled in the art. Gene presence, amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA, dot blotting (DNA or RNA analysis), or in situ hybridisation, using an appropriately labelled probe which may be based on a sequence

15   provided herein. Those skilled in the art will readily envisage how these methods may be modified, if desired.

We further describe cells containing the above-described nucleic acids/oligonucleotides. Such host cells such as prokaryote, yeast and higher eukaryote cells may be used for replicating DNA and producing the encoded polypeptide(s) protein.

20   Suitable prokaryotes include eubacteria, such as Gram-negative or Gram-positive organisms, such as *E. coli*, e.g. *E. coli* K-12 strains, DH5a and HB101, or Bacilli. Further hosts suitable for the above-described vectors include eukaryotic microbes such as filamentous fungi or yeast, e.g. Saccharomyces cerevisiae. Higher eukaryotic cells include insect and vertebrate cells, particularly mammalian cells including human cells or

25   nucleated cells from other multicellular organisms. In recent years propagation of vertebrate cells in culture (tissue culture) has become a routine procedure. Examples of useful mammalian host cell lines are epithelial or fibroblastic cell lines such as Chinese hamster ovary (CHO) cells, NIH 3T3 cells, HeLa cells or 293T cells. The host cells referred to in this disclosure comprise cells in *in vitro* culture as well as cells that are

30   within a host animal.

DNA may be stably incorporated into cells or may be transiently expressed using methods known in the art. Stably transfected mammalian cells may be prepared by transfecting cells with an expression vector having a selectable marker gene, and growing the transfected cells under conditions selective for cells expressing the marker gene. To

5   prepare transient transfectants, mammalian cells are transfected with a reporter gene to monitor transfection efficiency.

To produce such stably or transiently transfected cells, the cells should be transfected with a sufficient amount of the nucleic acid binding protein-encoding oligonucleotide to form the nucleic acid binding protein. The precise amounts of DNA

10   encoding the nucleic acid binding protein may be empirically determined and optimised for a particular cell and assay.

Host cells may be transfected or, preferably, transformed with expression or cloning vectors and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired

15   sequences. Heterologous DNA may be introduced into host cells by any method known in the art, such as transfection with a vector encoding a heterologous DNA by the calcium phosphate coprecipitation technique or by electroporation. Numerous methods of transfection are known to the skilled worker in the field. Successful transfection is generally recognised when any indication of the operation of this vector occurs in the host

20   cell. Transformation is achieved using standard techniques appropriate to the particular host cells used.

Incorporation of cloned DNA into a suitable expression vector, transfection of eukaryotic cells with a plasmid vector or a combination of plasmid vectors, each encoding one or more distinct genes or with linear DNA, and selection of transfected cells are well

25   known in the art (see, e.g. Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press).

Transfected or transformed cells are cultured using media and culturing methods
known in the art, preferably under conditions, whereby the nucleic acid binding protein
encoded by the DNA is expressed. The composition of suitable media is known to those in
the art, so that they can be readily prepared. Suitable culturing media are also

5   commercially available.

Oligonucleotides may be employed in a wide variety of applications, including
diagnostics and as research tools.

It is envisaged that the techniques described here may be usefully employed in the
production of libraries, such as expression libraries. In a preferred embodiment, the

10   techniques described here are applied to the production of oligonucleotides for
incorporation into phage display libraries. In a highly preferred embodiment, the
techniques described here are applied to the production of oligonucleotides comprised by
phage display libraries of nucleic acid binding proteins such as proteins comprising one or
more zinc finger motifs. However, it will be understood that the techniques described here

15   are of broad applicability and may be advantageously employed in the production of
selectively randomised oligonucleotides generally.

It is envisaged that the techniques described here may be advantageously combined
with or augmented by supplementary techniques. For example, if it was desired to
synthesise a position within an oligonucleotide using an NNN strategy, whilst retaining the

20   advantageous selective randomisation techniques, it would be understood that introducing
one or more NNN steps into the methods described here would not materially alter the
method(s) described here and would thus be encompassed.

Similarly, it would be possible to synthesise one or more codons of an
oligonucleotide using a trinucleotide step in the synthesis. Although this is likely to lead to

25   lower coupling efficiencies than the method(s) described here, and indeed trinucleotide
phosphoramidites are well known to be more difficult to manufacture than the
dinucleotide phosphoramidites described here if for some reason it was decided to

28

incorporate one or more such step(s) into the method(s) described here, a person skilled in the art would recognise that this would have no material effect on the methods.

Clearly, there is no requirement to limit the randomisation potential using the method(s) described here. The methods described herein may be advantageously applied to the generation of oligonucleotides which are randomised at any number or proportion of position(s) along their length, each position being independently constructed so as to potentially encode any of the twenty amino acids or a stop codon, or any subset thereof. Thus, use of the methods disclosed herein in the generation of a selectively randomised oligonucleotide comprising wholly randomised positions, or comprising wholly definite positions, falls within the scope of the present invention. Likewise, the methods described here may be usefully employed in the production of selectively randomised oligonucleotides for noncoding sequences.

Oligonucleotides are useful in biological, biochemical or chemical fields, for example those as described herein, as well as in related fields such as DNA 'machines' or 'computers', or applications involving the use of polymers or macromolecules such as nucleic acids or oligonucleotides for the storage and/or manipulation of data.

<u>EXAMPLES</u>

Example 1: Synthesis of Oligonucleotides

The synthesis of the dinucleotide blocks is essentially accomplished as described by Kumar et al. (Kumar and Poonian (1984) J. Org. Chem. 49 pp 4905-4912). All dimers have the expected ^{1}H and ^{31}P spectroscopic properties. Dimers are also available from Cruachem Ltd (Cruachem Ltd, Todd Campus, West of Scotland Science Park, Acre Road, Glasgow, G20 OUA Scotland, UK).

Automated oligonucleotide synthesis is performed on an Applied Biosystems 394 DNA/RNA synthesizer, using the standard 40 nmol scale synthesis protocols, with

coupling time of the dinucleotide units extended to 2 min. The synthesizer reagents are obtained from Applied Biosystems.

Mixtures of dinucleotide phosphoramidties are dried for 15 h over phosphorous pentoxide under vacuum, dissolved in acetonitrile (<10 p.p.m. of water, Labscan) and

5   attached to positions 5, 6, 7 and 8 of the DNA synthesizer. The oligonucleotides synthesized using the dinucleotide building blocks are treated with thiophenol (thiophenol/dioxane/triethylamine) to cleave the methyl phosphate ester protecting group, prior to ammonia treatment (16 h at 55°C). Volatile components are evaporated and the crude oligonucleotide is resuspended in H_2O. Oligonucleotides are used without further

10   purification. Optionally, oligonucleotides may be purified as discussed herein or using other purification techniques well known in the art.

Example 2: Generation Of Libraries Comprising Oligonucleotides

The cloning of oligonucleotides generated may be accomplished as follows:

Klenow polymerase and T4 DNA ligase are purchased from New England Biolabs

15   and used as recommended by the manufacturer. dNTPs are obtained from Boehringer Mannheim. Methods for plasmid purification, enzymatic reactions, cloning and bacterial transformation were performed as described (Sambrook et al (1989) molecular cloning - A laboratory manual – CSHL Press, Cold Spring Harbour, NY, USA). The oligonucleotide mixture is then subjected to a filling-in reaction with Klenow polymerase. The blunt-end

20   fragments thus produced are cloned in the *Eco*RV site of phagemid pBSks+.

The ligation mix is enriched for recombinant clones through *Eco*RV digestion prior to transformation in XL1-blue component bacterial cells. Recombinant clones are identified by colour screening on Xgal/IPTG/ampicillin plates. PCR amplification of the cloned sequence and gel electrophoresis analysis of the PCR products is used to analyse

25   the insert(s).

DNA sequencing is performed with dydeoxy terminator Taq cycle sequencing kit (Perkin Elmer) on an Applied Biosystems 373 automated DNA sequencer.

Example 3: Generation of Expression Libraries

Gene inserts for phage libraries may comprise a single oligonucleotide, or may be
5   constructed by end-to-end ligation of selectively randomised dsDNA 'minicassettes', made individually by annealing complementary template oligonucleotides (Fig. 1). In this Example, the latter is demonstrated. However, it will be plain to the skilled reader that the techniques described apply equally to libraries made using a single oligonucleotide.

The genes resulting from the end-to-end ligation are amplified by PCR and code
10   for zinc fingers in a suitable reading frame for cloning as fusions to the phage minor coat protein, pIII.

This Example uses the DNA-binding domain of the transcription factor Zif268 as a scaffold, since it contains three Cys_2-His_2 zinc fingers whose mode of binding is well understood.

15   In order to produce a selectively randomised α-helix of a zinc finger, the coding region is synthesised using DNA mini-cassettes (i.e.. selectively randomised oligonucleotides), such that helical positions –1 through 4 are encoded by one cassette (minicassette 2; Fig 1), while positions 4 through 6 are encoded by another cassette (minicassette 3; Fig 1). These double stranded 'cassettes' are synthesised with
20   complementary overhangs that anneal through the codon for the fourth α-helical residue, which is invariant.

Each 'cassette' actually comprises a library of oligonucleotides synthesised with appropriate codon randomisations so as to code for a given subset of amino acids.

31

Figure 1A shows that a 'smart library' of zinc finger genes can be created using three 'mini-cassettes': the first cassette is a single sequence and codes for the invariant β-sheet region, while the second and third cassettes contain randomisations of the α-helix. Fig 1B shows that each of the 'library mini-cassettes' comprises numerous

5   oligonucleotides created through a limited number of solid-phase syntheses: minicassette 2 requires oligonucleotides from 12 pairs of syntheses, while minicassette 3 requires oligonucleotides from three pairs of syntheses. Each oligonucleotide synthesis is designed to introduce a very limited variability into each cassette – the library complexity is increased by the use of oligonucleotides from multiple syntheses and by the combination

10   of the two mini-cassettes.

The library is constructed according to the following protocol; Single-stranded template oligonucleotides are phosphorylated in a kinase reaction prior to assembly (100 pmol of each oligonucleotide in 10 μl of 1 × T4 kinase buffer, containing 1 mM dATP and 10 U T4 polynucleotide kinase, 37°, 1 hr). Complementary single-stranded template

15   oligonucleotides are annealed pairwise to form double-stranded minicassettes (Fig. 1): 100 pmol of each oligonucleotide (or, for smart randomisation, 100 pmol of each strand mixture) are mixed in 1 × T4 ligase or kinase buffer, to a final DNA concentration of 10 pmol/μl. Annealing is by heating to 94° and then cooling slowly (~1 hr) to room temperature. The resulting dsDNA minicassettes are combined and ligated by adding an

20   equal volume of 1 × T4 ligase buffer and 8 μl (3200 U) of T4 ligase per 100 μl (16°, 20 hr).

Full-length genes are amplified by PCR from the ligation mixture with primers that introduce *Not*I and *Sfi*I restriction sites for cloning into phage vector Fd-TET-SN(Y. Choo and A. Klug, *Proc. Natl. Acad. Sci. U.S.A.* 91, 11163 (1994)).

25   Thorough digestion with these endonucleases facilitates high-efficiency ligation into similarly prepared phage vector (200 U enzyme per 40 μg DNA, with 8 hr incubation in appropriate temperatures and buffers, adding enzymes in stages at 2-hr intervals). Typically, 1 μg of pure phage vector is ligated with a 5-fold excess of gene cassette insert

(1 × T4 ligase buffer, 3 μl T4 ligase, 30 μl total volume, 16°, 20 hr). Ligation reactions are prepared for electroporation by washing twice in an equal volume of chloroform and precipitating by adding 1/10 volume sodium acetate (pH 5.5) and 3 volumes of ethanol. DNA pellets are washed with 70% ethanol and resuspended in sterile water to a final

5   concentration of 200 ng/μl.

The phage library is cloned by electroporation of recombinant vector into a suitable strain of *E. coli*, such as TG1. Typically, 0.5 μg of recombinant phage vector can be used with 100 μl of electrocompetent cells (W. J. Dower, J. F. Miller and C. W. Ragsdale, *Nucleic acids Research* 16, 6127 (1988)), yielding up to ~10^6 library

10   transformants (2 mm path cuvette, 2.5 kV, 25 μF, 200 ohms). After pulsing, cells are immediately resuspended in 1 ml SOC and incubated without shaking (37°, 1 hr). Fd-TET-SN confers tetracycline resistance allowing positive selection of bacterial transformants by plating on 2 × YT-agar plates, containing 15 μg/ml tetracycline (37°, 16 hr).

15   Selection from libraries is accomplished according to the following protocol; Phage are prepared for selections by scraping library transformant colonies into 2 × YT liquid medium, containing 15 μg/ml tetracycline and 50 μM zinc chloride, and incubating in an orbital shaker (30°, 220 rpm, 16 hr). The culture supernatant containing phage particles is collected by centrifugation (3700 g, 15 min). For each selection, the

20   supernatant is diluted 1:10 in 1 ml PBS containing 1% (w/v) Marvel, 1% (v/v) Tween-20 and 20μg/ml sonicated salmon sperm DNA. The phage mixtures are added to streptavidin-coated tubes or wells (Roche) which have been pre-coated with biotinylated target DNA (made by annealing two complementary oligonucleotides, one of which is biotinylated).

The selection procedure described in this Example is for use with streptavidin-

25   coated tubes and a total reaction volume of 1ml, but by scaling down to a 200μl volume the process can easily be adapted to a 96-well microtitre plate format.

1 pmol target DNA is coated on each tube, in 50 µl PBS/Zn (~20°, 15 min). The addition of 1 ml 4% (w/v) Marvel blocking agent helps to reduce non-specific binding. After blocking (~20°, 1 hr) tubes are emptied, re-filled with 1 ml phage binding mixtures, and left to equilibrate (~20°, 1 hr). Washing steps are then carried out to remove all

5   unbound phage (20 washes with 1 ml PBS/Zn containing 2% (w/v) Marvel, 1% (v/v) Tween-20, followed by one wash with PBS/Zn alone). Retained phage are eluted in 100 µl 0.1 M triethylamine, removed to a separate container and immediately neutralised with an equal volume of 1 M Tris-HCl, pH 7.4. Eluted phage can be stored at -20°.

10   50µl of the eluted phage are used to infect 0.3 ml of a logarithmic-phase culture of *E. coli* TG1. The bacteria are derived from colonies grown freshly on M9 minimal agar as this ensures expression of the F' pilus, which facilitates phage infection. Bacteria are infected by the addition of phage and incubating without shaking (37, 1 hr). Bacteria are then transferred to 2 - 5 ml of 2 × YT, containing 15 µg/ml tetracycline and 50 µM zinc chloride and grown, as before, to prepare phage supernatant (30°, 220 rpm, 16 hr).

15   Subsequent rounds of selection are carried out as described above. The amount of competitor DNA may optionally be increased in later rounds to increase the stringency of selection. The progress of individual selections may be monitored by plating out infections to estimate phage yield after each round of selection. 3 - 5 rounds of selection are usually sufficient to enrich target-binding clones.

20   Phage yield is estimated as follows: Phage titre from selection eluates and culture supernatants is estimated by using 1 µl of phage sample to infect 1 ml of a logarithmic-phase culture of *E. coli* (37°, 1 hr, no shaking). Infections are serially diluted ten-fold with individual dilutions being spread on 2 × YT-agar plates containing 15 µg/ml tetracycline. After incubation (37°, 16 hr), individual colonies are counted to give an indication of the

25   colony forming units (phage titre) in the original sample. ·

After selections, pools of complementary phage can be recombined directly, or grown up to quantities for further study or use, for example-individual clones may be tested for binding to their respective 5 bp target sites using ELISA.

Each of the applications and patents mentioned above, and each document cited or referenced in each of the foregoing applications and patents, including during the prosecution of each of the foregoing applications and patents ("application cited documents") and any manufacturer's instructions or catalogues for any products cited or

5   mentioned in each of the foregoing applications and patents and in any of the application cited documents, are hereby incorporated herein by reference. Furthermore, all documents cited in this text, and all documents cited or referenced in documents cited in this text, and any manufacturer's instructions or catalogues for any products cited or mentioned in this text, are hereby incorporated herein by reference. In particular, we hereby incorporate by

10 reference International Patent Application Numbers PCT/GB00/02080, PCT/GB00/02071, PCT/GB00/03765, United Kingdom Patent Application Numbers GB0001582.6, GB0001578.4, and GB9912635.1 as well as US09/478513.

Various modifications and variations of the described methods and system described here will be apparent to those skilled in the art without departing from the scope

15 and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following

20 claims.

CLAIMS

1.   A method for making a synthetic oligonucleotide comprising

(i) coupling a dinucleotide phosphoramidite to the 3' position of a nucleotide base;

(ii) coupling a mononucleotide phosphoramidite to the oligonucleotide of (i); and

5   (iii) repeating steps (i) and (ii) until the desired length oligonucleotide is produced.

2.   A method according to Claim 1, in which the oligonucleotide is synthesised on a solid support.

3.   A method for making a selectively randomised synthetic oligonucleotide comprising:

10   (i) providing a starting material coupled to a suitable support in a nucleic acid synthesiser;

(ii) deprotecting the starting material at the 3' position;

(iii) coupling a dinucleotide phosphoramidite to the 3' position ;

(iv) deprotecting the new 3' position of the extended oligonucleotide chain;

15   (v) coupling a mononucleotide phosphoramidite to the 3' position; and

(vi) repeating steps (ii) to (v) until the desired length oligonucleotide is produced.

4.   A method according to Claim 3, in which at each repetition of step (iii) a specific dinucleotide phosphoramidite or specific mixture of dinucleotide phosphoramidite(s) is used in the coupling.

5.   A method according to Claim 3 or 4, in which at each repetition of step (iii) a dinucleotide phosphoramidite or mixture of dinucleotide phosphoramidite(s) selected from the following table is incorporated into the growing oligonucleotide molecule:

Nucleotide		Optimal		Non-optimal	
		human codon	(codes for)	codon	(codes for)
Di-	Mono				
XY	Z	XY-Z	(aa)	XY-Z	(aa)
AA	G	AA-G	Lys	N/A	
AA	C	AA-C	Asn	N/A	
AT	G	AT-G	Met	N/A	
AT	C	AT-C	Ile	N/A	
CA	G	CA-G	Gln	N/A	
CA	C	CA-C	His	N/A	
GA	G	GA-G	Glu	N/A	
GA	C	GA-C	Asp	N/A	
TG	G	TG-G	Trp	N/A	
TG	C	TG-C	Cys	N/A	
AC	C	AC-C	Thr	AC-G	Thr
CC	C	CC-C	Pro	CC-G	Pro
CG	C	CG-C	Arg	CG-G	Arg
CT	C	CT-C	Leu	CT-G	Leu
GC	C	GC-C	Ala	GC-G	Ala
GG	C	GG-C	Gly	GG-G	Gly
GT	C	GT-C	Val	GT-G	Val
TC	C	TC-C	Ser	TC-G	Ser
TA	C	TA-C	Tyr	TA-G	STOP
TT	C	TT-C	Phe	TT-G	Leu

6.   A method according to any preceding claim, in which the mononucleotide

5   phosphoramidite comprises a mixture of G/C mononucleotide phosphoramidites.

7.   A method according to Claim 6, in which the mononucleotide phosphoramidite comprises a 50:50 ratio of G:C mononucleotide phosphoramidites.

8.   A method according to any of Claims 1 to 5, in which the mononucleotide phosphoramidite comprises a mixture of A/T mononucleotide phosphoramidites.

5   9.   A method according to Claim 8, in which the mononucleotide phosphoramidite comprises a 50:50 ratio of A:T mononucleotide phosphoramidites.

10.   A method according to any preceding claim, in which the oligonucleotide comprises an open reading frame (ORF).

11.   A method according to Claim 10, in which the ORF comprises an optimal human
10   codon, preferably selected from the optimal human codons as shown below:

Nucleotide		Optimal		Non-optimal	
Di-	Mono	human codon	(codes for)	codon	(codes for)
XY	Z	XY-Z	(aa)	XY-Z	(aa)
AA	G	AA-G	Lys	N/A	
AA	C	AA-C	Asn	N/A	
AT	G	AT-G	Met	N/A	
AT	C	AT-C	Ile	N/A	
CA	G	CA-G	Gln	N/A	
CA	C	CA-C	His	N/A	
GA	G	GA-G	Glu	N/A	
GA	C	GA-C	Asp	N/A	
TG	G	TG-G	Trp	N/A	
TG	C	TG-C	Cys	N/A	
AC	C	AC-C	Thr	AC-G	Thr
CC	C	CC-C	Pro	CC-G	Pro

CG	C	CG-C	Arg	CG-G	Arg
CT	C	CT-C	Leu	CT-G	Leu
GC	C	GC-C	Ala	GC-G	Ala
GG	C	GG-C	Gly	GG-G	Gly
GT	C	GT-C	Val	GT-G	Val
TC	C	TC-C	Ser	TC-G	Ser
TA	C	TA-C	Tyr	TA-G	STOP
TT	C	TT-C	Phe	TT-G	Leu

12. A method according to Claim 10 or 11, in which codons of the ORF are synthesised from dinucleotide phosphoramidites (XY) and mononucleotide phosphoramidites (Z) in the XY-Z conformation such that the dinucleotide phosphoramidite forms the first two bases of the codon, the mono nucleotide phosphoramidite forming the last base of the codon.

13. A method according to any of Claims 10 to 12, in which the oligonucleotide encodes one or more zinc finger motif(s) or part(s) thereof.

14. A method according to Claim 13, in which the oligonucleotide is selectively randomised at positions other than the conserved CYS-HIS motif(s).

15. An oligonucleotide produced by a method according to any preceding claim.

16. A library comprising an oligonucleotide produced according to any of Claims 1 to 14.

17. A library according to Claim 16 in which the library is a phage display library.

1/1

A

| β | β | α-helix | linker |

-1 1 2 3 4   4 5 6

Minicassette1   Minicassette2   Minicassette3
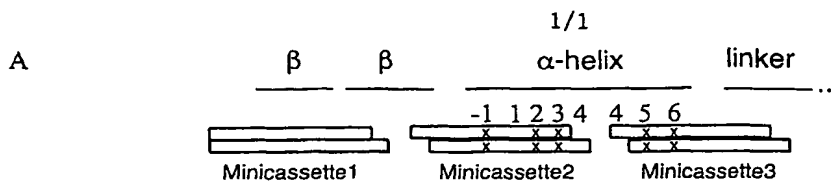
B   Minicassette 1

5'-GCGGAAGAGAGGCCCTACGCATGCCCTGTCGAGTCCTGCGATCGC
   CGCCTTCTCTCCGGGATGCGTACGGGACAGCTCAGGACGCTAGCGGCGA-5'

Minicassette 2

```
     Fl  -1  +1 +2 +3
CGCTTTTCTxxxTCGxxxxx        -TOP
      AAAGAxxxAGCxxxxxGGAAT  -BOTTOM
```

Code for:

TOP-	BOTTOM-	-1	+2	+3
CGCTTTTCTcrsTCGgmcca	TAAGGtggkcCGAsygAGAAA	R,Q,H	D,A	H
CGCTTTTCTcrsTCGgmcav	TAAGGbtgkcCGAsygAGAAA	R,Q,H	D,A	N,S,T
CGCTTTTCTcrsTCGmcgh	TAAGGdcgkcCGAsygAGAAA	R,Q,H	D,A	V,A,D
CGCTTTTCTcrsTCGmrsca	TAAGGtgsykCGAsygAGAAA	R,Q,H	RQHKSN	H
CGCTTTTCTcrsTCGmrsav	TAAGGbtsykCGAsygAGAAA	R,Q,H	RQHKSN	N,S,T
CGCTTTTCTcrsTCGmrsgh	TAAGGdcsykCGAsygAGAAA	R,Q,H	RQHKSN	V,A,D
CGCTTTTCTrmcTCGgmcca	TAAGGtggkcCGAgkyAGAAA	N,D,A,T	D,A	H
CGCTTTTCTrmcTCGgmcav	TAAGGbtgkcCGAgkyAGAAA	N,D,A,T	D,A	N,S,T
CGCTTTTCTrmcTCGgmcgh	TAAGGdcgkcCGAgkyAGAAA	N,D,A,T	D,A	V,A,D
CGCTTTTCTrmcTCGmrsca	TAAGGtgsykCGAgkyAGAAA	N,D,A,T	RQHKSN	H
CGCTTTTCTrmcTCGmrsav	TAAGGbtsykCGAgkyAGAAA	N,D,A,T	RQHKSN	N,S,T
CGCTTTTCTrmcTCGmrsgh	TAAGGdcsykCGAgkyAGAAA	N,D,A,T	RQHKSN	V,A,D

Minicassette 3

```
   Fl +4  +5 +6
5'- CCTTAxxxxxCATATCCGCATCCACA    -TOP
       xxxxxGTATAGGCGTAGGTGTGGCC  -BOTTOM
```

Code for:

TOP-	BOTTOM-	+5	+6
CCTTayccrgCATATCCGCATCCACA	CCGGTGTGGATGCGGATATGcyggr	I,T	R,Q
CCTTaycghaCATATCCGCATCCACA	CCGGTGTGGATGCGGATATGtdcgr	I,T	V,A,E
CCTTaycamsCATATCCGCATCCACA	CCGGTGTGGATGCGGATATGsktgr	I,T	K,N,T

Key to Randomised Nucleotides

m=A/C	s=G/C	b=C/G/T	v=A/C/G
r=A/G	y=T/C	n=A/C/G/T	h=A/C/T
w=A/T	k=T/G		d=A/G/T

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7   C07H21/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED
Minimum documentation searched (classification system followed by classification symbols)
IPC 7   C07H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	ONO A ET AL: "THE SYNTHESIS OF BLOCKED TRIPLET-PHOSPHORAMIDITES AND THEIR USE IN MUTAGENESIS" NUCLEIC ACIDS RESEARCH, OXFORD UNIVERSITY PRESS, SURREY, GB, vol. 23, no. 22, 1995, pages 4677-4682, XP001026095 ISSN: 0305-1048 cited in the application whole document -/--	1-17

[X] Further documents are listed in the continuation of box C.   [ ] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
29 January 2002	07/02/2002

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Klein, D

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category°	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
Y	CHAIX C ET AL: "3'-3'-linked oligonucleotides: synthesis and stability studies" BIOORGANIC & MEDICINAL CHEMISTRY LETTERS, OXFORD, GB, vol. 6, no. 7, 9 April 1996 (1996-04-09), pages 827-832, XP004134966 ISSN: 0960-894X see page 828 bottom and scheme 1 upper part and left part.	1-17